

University of Groningen

Genomic distribution of the small multidrug resistance protein EmrE over 29 Escherichia coli strains reveals two forms of the protein

Kolbusz, Magdalena A.; Slotboom, Dirk J.; Lolkema, Juke S.

Published in:
Febs Journal

DOI:
[10.1111/febs.12065](https://doi.org/10.1111/febs.12065)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2013

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kolbusz, M. A., Slotboom, D. J., & Lolkema, J. S. (2013). Genomic distribution of the small multidrug resistance protein EmrE over 29 Escherichia coli strains reveals two forms of the protein. *Febs Journal*, 280(1), 244-255. <https://doi.org/10.1111/febs.12065>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Genomic distribution of the small multidrug resistance protein EmrE over 29 *Escherichia coli* strains reveals two forms of the protein

Magdalena A. Kolbusz¹, Dirk J. Slotboom² and Juke S. Lolkema¹

¹ Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands

² Membrane Enzymology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands

Keywords

antiparallel dimers; EmrE; membrane protein orientation; multidrug resistance; SMR family

Correspondence

J. S. Lolkema, Molecular Microbiology, Nijenborgh 7, 9747AG Groningen, The Netherlands
Fax: +31 50 363 2154
Tel: +31 50 363 2155
E-mail: j.s.lolkema@rug.nl

(Received 16 July 2012, revised 8 November 2012, accepted 12 November 2012)

doi:10.1111/febs.12065

Analysis of the genomes of 29 *Escherichia coli* strains revealed two different versions of the EmrE protein, a member of the small multidrug resistance family. The versions are different in length and contain 110 residues (EMRE₁₁₀) and 165 residues (EMRE₁₆₅). The N-terminal extension found in the longer sequence has the properties of a signal sequence, i.e. contains at the extreme N-terminus a hydrophobic region followed by a predicted cleavage site. Analysis of the genetic context of the genes in the different strains showed that all of the genes encoding EMRE₁₆₅ had the same context, whereas the genes encoding EMRE₁₁₀ were distributed over four different, but similar, contexts. The different genetic contexts corresponded to the branching of the phylogenetic tree of the *emrE* genes. Membrane topology studies using translational fusions with the two reporter proteins alkaline phosphatase and green fluorescent protein showed the well-described dual topology mode of insertion of EMRE₁₁₀. In contrast, but in line with the presence of the signal sequence, EMRE₁₆₅ was inserted in a single orientation into the membrane, with the C-terminus in the periplasm. The N-terminal region was removed from the protein after insertion into the membrane. In contrast to cells expressing EMRE₁₁₀, cells expressing only mature EMRE₁₆₅ were not able to grow on plates containing ethidium bromide. The results suggest that if dimers were formed from EMRE₁₆₅ monomers with the same orientation in the membrane, they would not be active in drug extrusion.

Introduction

Members of the small multidrug resistance (SMR) family are secondary transporters found exclusively in prokaryotes. The transporters extrude toxic compounds such as ethidium bromide (EtBr), methyl viologen and tetraphenyl phosphonium from the cells, and also render the cells resistant to antibiotics such as cephalosporins [1,2], aminoglycosides [3,4], and β -lactams [5–7]. The extrusion process is energized by the electrochemical proton gradient across the membrane

[8]. The proteins are the smallest secondary efflux proteins known to date, and consist of 100–140 residues that fold into four tightly packed α -helices [9]. The catalytic unit is a dimer. The best studied protein from the family is EmrE of *Escherichia coli*, which consists of 110 residues. EmrE is a so-called dual topology protein; that is, it inserts into the membrane in two opposite orientations [10–13]. The presence of both orientations in the membrane allows the formation of

Abbreviations

AP, alkaline phosphatase; BAD, biotin acceptor domain; EtBr, ethidium bromide; GFP, green fluorescent protein; HGT, horizontal gene transfer; LIC, ligase-independent cloning; SMR, small multidrug resistance; TMS, transmembrane segment.

two types of dimer: those in which two monomers of the same orientation interact (parallel), and those in which the subunits in the dimer have opposite orientation (antiparallel). Evidence in favor of the antiparallel dimer [14–18], the parallel dimer [19,20] or both [21,22] has been presented. Thus, the issue is still under debate.

The EmrE used in these studies was taken from *E. coli* strain K-12. Surprisingly, an analysis presented here of 29 strains of *E. coli* for which the complete genome sequence was available showed that 10 strains encoded a longer version of EmrE consisting of 165–170 residues (EMRE₁₆₅), rather than the 110 residues of the K-12 variant (EMRE₁₁₀). The longer version of EmrE was extended at the N-terminus, and contained an additional hydrophobic segment that was predicted to be part of a signal sequence (SIGNALP) [23,24]. Consequently, the long version might be directed in one orientation into the membrane, which, if it was capable of forming dimers, would allow only the formation of the parallel dimers.

This study focuses on the properties of the long version of EmrE in comparison with the short version. Genomic localization of genes encoding EmrE proteins demonstrates that the two versions are found in different loci. The reporter proteins alkaline phosphatase (AP) and green fluorescent protein (GFP) fused to the C-terminus confirmed that EMRE₁₆₅ inserts into the membrane in a single orientation, and processing of the signal sequence is demonstrated by the loss of the additional hydrophobic segment. Finally, the activity of matured EMRE₁₆₅ was determined by checking the ability of cells expressing EMRE₁₆₅ to grow in the presence of EtBr.

Results

Distribution of EmrE over 29 strains of *E. coli*

Analysis of the genome of 29 *E. coli* strains available from the National Center for Biotechnology Information (NCBI) database showed that 18 strains contained a single copy of the *emrE* gene, three strains contained two copies, and eight strains lacked the gene (Table 1). Analysis of the genomic context divided the 24 *emrE* genes into two groups. A group of 10 genes showed a well-defined context. At one side, they were flanked by flagellar region IIIB, containing genes involved in flagellum assembly, and at the other side by a putative kinase inhibitor gene and a regulatory gene (Fig. 1; FLA). The other 14 *emrE* genes were flanked downstream by a gene encoding a phage recombinase and upstream by a gene encoding the REN protein.

Table 1. Distribution of EmrE proteins over *E. coli* strains. EmrE (GI): GI number NCBI in protein database. Insertion site: see text and Fig. 1.

Strain	EmrE (GI)	Insertion site	Number of residues
ED1a	218689931	FLA	165
UT189	91211154	FLA	165
S88	218558793	FLA	165
536	110642042	FLA	165
APEC01	117624076	FLA	165
11128	260868529	FLA	165
HS	157161407	FLA	165
CFT073	161486199	FLA	165
55989	218695541	FLA	165
	218694205	REN_4	110
SE11	209919358	FLA	165
	209917769	REN_4	110
E24377A	157158420	REN_4	110
CB9615	291281465	REN_4	110
EC4115	209400746	REN_4	110
	209397090	REN_1	110
TW14359	254792602	REN_1	110
O157:H7	15801338	REN_1	110
Sakai	15830868	REN_1	110
BW2952	238899820	REN_2	110
K-12_DH10B	170080124	REN_2	110
K-12_MG1655	16128526	REN_2	110
REL606	254160610	REN_3	110
BL21-GOLD	253774471	REN_3	110
IAI1	–	–	–
IAI39	–	–	–
12009	–	–	–
SMS-3-5	–	–	–
UMN026	–	–	–
ATCC8739	–	–	–
E2348/69	–	–	–
11368	–	–	–

Differences in the upstream region of the *REN* gene discriminated four different subgroups (Fig. 1; REN_1 to REN_4). The variable upstream regions contained many genes encoding integrases (REN_1, REN_2, and REN_3), replication proteins (REN_1 and REN_4), exonucleases (REN_2), and IS elements (REN_2) that may be involved in the insertion/deinsertion of mobile elements, suggesting that the *emrE* gene, together with the two flanking genes, was inserted in these sites at some point in evolution. The two *emrE* genes found in the genome of strain EC4115 are both of the REN type, whereas strains 55989 and SE11 contain both an FLA type and a REN type.

A phylogenetic tree of the *emrE* genes found in the different *E. coli* strains was constructed on the basis of the alignment of the corresponding parts of the DNA (see below). The alignment contained no gaps, and revealed pairwise nucleotide sequence identities of

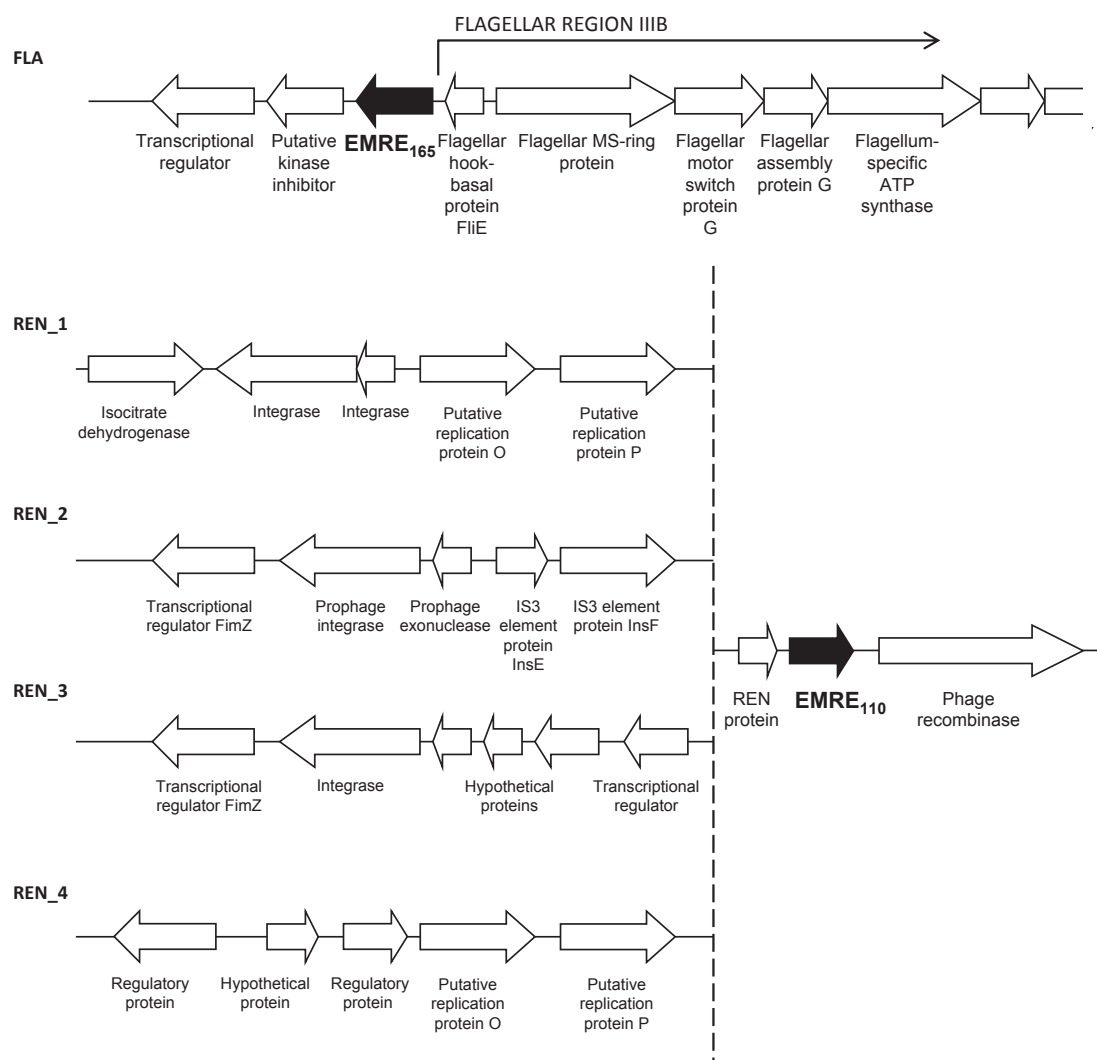


Fig. 1. Genetic context of long and short versions of EmrE. See text for explanation.

between 90% and 99%, indicating that the tree represents the very recent divergence of the genes in the different strains. The tree revealed two well-separated clusters represented by the FLA and REN types of gene (Fig. 2). With the exception of the gene in strain 55989 (GI: 218694205), the different genomic contexts of the REN types were found as different branches on the tree as well, indicating that the divergence of the genes took place after insertion into the genome. No correlation was found between the phylogenetic tree of the genes encoding EmrE proteins and the whole-genome phylogenetic tree of *E. coli* strains computed with feature frequency profiles [23]. The strains containing the FLA-type EmrE did not group together on one branch of the strain tree, and, similarly, the strains containing the REN type were

scattered over the tree. This suggests that the genes encoding the different types of EmrE were spread by horizontal gene transfer (HGT).

The well-studied EmrE from strain K-12 consists of the usual 110 residues found for most of the members of the SMR family. The K-12 protein is of the REN type, and all EmrE proteins of this type from the other strains were annotated in the databases as proteins of the same size. Surprisingly, FLA-type EmrE proteins were annotated as proteins containing 165 or 170 residues. The additional base pairs resulting in the larger proteins were found at the 5'-end of the gene. Inspection of the nucleotide sequence revealed that all FLA-type EmrE proteins had the same upstream sequence, which differed from that of the REN-type EmrE proteins. Annotation differences (165 versus 170) were

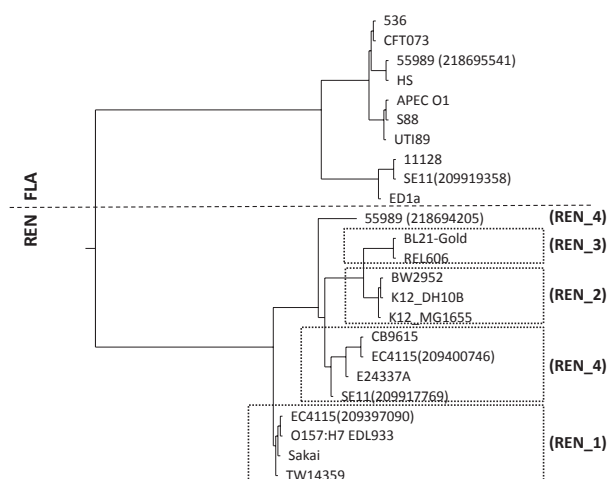


Fig. 2. Phylogenetic tree of the *emrE* genes from different *E. coli* strains. The tree was constructed on the basis of a multiple sequence alignment of the corresponding 330-nucleotide fragments of the genes. Sequences encoding for EMRE₁₆₅ were truncated at the 5'-end, resulting in nucleotides with the same length as those encoding EMRE₁₁₀. This avoided a bias in the tree because of the different lengths of the genes encoding EMRE₁₆₅. The insertion sites FLA and REN_1–4 are defined in Fig. 1. EmrE proteins are indicated by the strain in which they were identified.

mainly the consequence of two closely spaced initiation codons. It follows that the genes in the FLA and REN insertion sites encode a long and a short version of the EmrE protein, EMRE₁₆₅ and EMRE₁₁₀, respectively.

Sequence analysis of EMRE₁₁₀ and EMRE₁₆₅

The difference between the nucleotide sequences encoding EMRE₁₆₅ and EMRE₁₁₀ is in the 5'-end of the genes (Fig. 3A). From a position that is 76 nucleotides upstream of the start codon of the short version, the two sequences overlap and are highly identical, i.e. 91% in the 536 and K-12 strains. The corresponding parts contain identical ribosomal binding sites, and ATG start and UAA stop codons. Upstream of the corresponding parts, the sequences are unrelated. EMRE₁₆₅-encoding DNA contains a GTG start codon that is in frame with the ATG start codon of the short version, whereas this is not the case in EMRE₁₁₀-encoding DNA. It follows that the messenger produced from EMRE₁₆₅-encoding DNA is likely to produce a mixture of EMRE₁₆₅ and EMRE₁₁₀.

The widely used membrane topology predictor TMHMM 2.0 [24] predicts the presence of four trans-membrane segments (TMSs) in EMRE₁₁₀, which is in agreement with experimental data [25–27]. The protein is predicted to have no preference for one of the two

orientations in the membrane, because of the lack of a positive charge bias over the loops at both sides of the membrane (positive-inside rule) [13]. The prediction is in line with the dual topology character of EMRE₁₁₀, which was convincingly demonstrated experimentally (Fig. 3B) [10,11,13,28]. In the EMRE₁₆₅ sequence, TMHMM predicts an additional TMS positioned at the N-terminus. The putative TMS would be connected to the four TMS bundle by a loop of 34 residues. The signal peptide predictor SIGNALP [29,30] predicted the N-terminal TMS to be part of a signal sequence with maximal cleavage site probability between positions 23 and 24. The presence of the signal sequence suggested that the protein would be inserted in the membrane in one specific orientation and, following cleavage of the leader sequence by leader peptidase, would leave the matured protein in the N_{out}C_{out} orientation (Fig. 3B).

Membrane topology of EMRE₁₁₀ and EMRE₁₆₅

The ORFs encoding EMRE₁₁₀ and EMRE₁₆₅ of strains K-12 and 536, respectively, were cloned in pLIC vectors [31], yielding pLIC_EMRE₁₁₀ and pLIC_EMRE₁₆₅, respectively. The pLIC vectors produce the proteins with a His₆-tag at the N-terminus and AP (pLIC1) or GFP (pLIC2) fused to the C-terminus. The two reporters AP and GFP allow for the determination of the cellular location of the C-termini of EMRE₁₁₀ and EMRE₁₆₅. GFP is properly folded and fluorescent only when targeted to the cytoplasm, whereas AP is enzymatically active only when exported to the periplasm. High GFP fluorescence and low AP activity indicate that the C-terminus is located in the cytoplasm, whereas high AP activity and low GFP fluorescence indicate a periplasmic C-terminus localization. Significant activity of both reporters indicates dual topology. It is important to stress here that expression of the EMRE₁₆₅-encoding gene from pLIC plasmids is likely to produce reporter fusions of both EMRE₁₆₅ and EMRE₁₁₀, only the former of which contains the N-terminal His₆-tag. To produce only EMRE₁₆₅, the Met at position 56 (start of EMRE₁₁₀) was mutated to Ala, yielding the vectors pLIC1_EMRE₁₆₅(M56A) and pLIC2_EMRE₁₆₅(M56A). The same set of plasmids was constructed by using the pBAD-cLIC vectors that produce the fusion proteins with a His₁₀-tag at the C-termini of the reporter proteins. In this case, both long and short versions produced from the EMRE₁₆₅-encoding gene contain the His-tag (see Table 2 for an overview of the constructs).

The normalized activities (see Experimental procedures) of the reporter proteins GFP and AP fused to EMRE₁₁₀ produced from both pLIC_EMRE₁₁₀ and

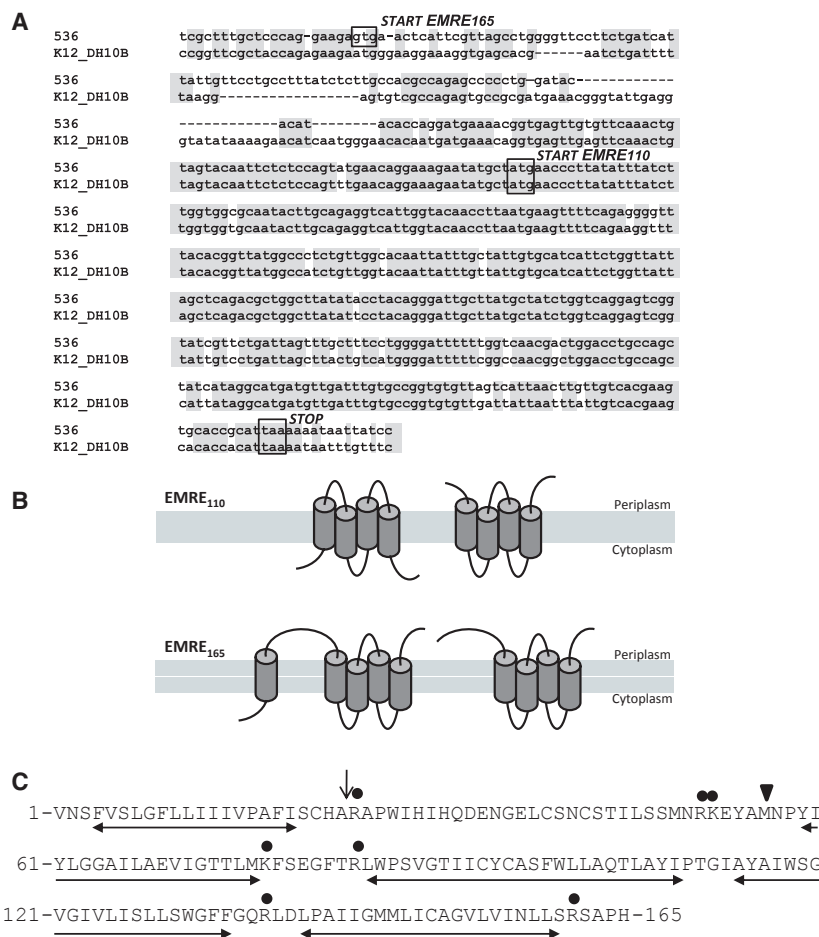


Fig. 3. Comparison of the long and short versions of EmrE. (A) Alignment of base sequences containing the genes encoding EMRE₁₆₅ from strain 536 and EMRE₁₁₀ from strain K-12. The alignment is annotated with start and stop codons. (B) Dual topology of EMRE₁₁₀ (top) and topology model for EMRE₁₆₅ before (bottom, left) and after (bottom, right) cleavage of the signal sequence. The signal sequence was predicted by the SIGNALP server with a probability of 0.967 (HMM mode). (C) Amino acid sequence of the EMRE₁₆₅ preprotein. Double-headed arrows indicate the positions of predicted TMSs. The putative signal cleavage site is indicated by an arrow, the first residue of EMRE₁₁₀ by a triangle, and positively charged residues by dots.

pBADcLIC_EMRE₁₁₀ were similar (Fig. 4A,B), indicating that the short version inserted into the membrane was distributed more or less equally over the two orientations (dual topology), as documented many times before [10,11,13]. The distribution shifted significantly to the orientation with the C-terminus in the periplasm with the EMRE₁₆₅-encoding gene in plasmids pLIC_EMRE₁₆₅ and pBADcLIC_EMRE₁₆₅ (Fig. 4A,B). Importantly, the signals obtained from the reporters are likely to be the sum of the contributions of both long and short EmrE versions. Apparently, the long version contributes significantly to the fraction of molecules with the C-terminus in the periplasm. Within the limits of experimental error, the result was independent of the position of the His-tag at the N-terminus or C-terminus of the fusion proteins

(Fig. 4A,B). Also, the levels of expression deduced from the reporter activities were of the same order of magnitude for the different constructs. In contrast, the level of expression dropped significantly for both pLIC_EMRE₁₆₅(M56A) and pBADcLIC_EMRE₁₆₅(M56A), when only the long version was produced. Relative to EMRE₁₁₀, the distribution of the version with the His-tag at the N-terminus was shifted to the orientation with the C-terminus in the periplasm, but a significant fraction had the opposite orientation (Fig. 4A). With the His-tag at the C-terminus, the orientation of EMRE₁₆₅(M56A) with the C-terminus in the periplasm was dominant (Fig. 4B). Optimization of the expression level of the latter construct by using a range of inducer concentrations resulted in a two-fold increase in expression level. Importantly, the

Table 2. Vectors used in this study to express EMRE₁₁₀, EMRE₁₆₅, and EMRE₁₆₅(M56A). All vectors are pBAD24-based, and use the arabinose promoter for induction of expression. N-terminus: tag fused at the N-terminus of EmrE. C-terminus: tag/reporter fused at the C-terminus of EmrE.

	N-terminus	C-terminus	Reference
LIC vectors			
pLIC1	His ₆ -tag	AP	31
pLIC2	His ₆ -tag	GFP	31
pBADcLIC-AP	–	AP/His ₁₀ -tag	This study
pBADcLIC-GFP	–	GFP/His ₁₀ -tag	39
pBAD24 vectors			
pBAD24	–	–	Invitrogen
pBADhis	His ₆ -tag	–	40
pBADBAD	His ₆ -tag/BAD	–	41

orientation of the protein in the membrane was unaffected (Fig. 4B).

The EmrE–GFP fusion proteins produced from the pBADcLIC–GFP and pLIC2 vectors were purified from isolated membranes by Ni²⁺–nitrilotriacetic acid affinity chromatography. The yield of the proteins was too low to be visualized by protein staining after SDS/PAGE or western blotting with antibodies raised against GFP or the His-tag. Rather, the proteins were visualized by in-gel GFP fluorescence (Fig. 5). The EMRE₁₁₀ fusion protein expressed from the pBADcLIC–GFP vector and carrying the His₁₀-tag at the C-terminus bound strongly to the resin, and was only observed in the elution fractions. The band represents the membrane-bound protein molecules that originally had the C-terminus in the cytoplasm, where GFP matures to its fluorescent state. The same EMRE₁₁₀–GFP fusion was observed with membranes isolated from cells containing EMRE₁₆₅-encoding DNA from

which both EMRE₁₁₀ and EMRE₁₆₅ are produced. In contrast, no fluorescent EMRE₁₆₅–GFP fusion protein was observed, suggesting that the GFP moiety was exported to the periplasm during biosynthesis. In agreement with this, no fluorescent band was observed when EMRE₁₆₅ was produced alone [EMRE₁₆₅(M56A)] (Fig. 5A). Apparently, within the detection limit, all EMRE₁₆₅–GFP fusion proteins insert with the C-terminus in the periplasm, which is in agreement with the orientation assay above (Fig. 4B). The majority of the EMRE₁₁₀–GFP fusion protein carrying the His₆-tag at the N-terminus produced from the pLIC2 vector was found in the elution fraction, but binding to the resin was clearly weaker, leaving significant fractions in the flowthrough and wash steps (Fig. 5B). As expected, EMRE₁₁₀ produced from EMRE₁₆₅-encoding DNA showed up in the flowthrough because it was translated from an internal ORF without a His-tag. No clear band for EMRE₁₆₅ was observed. In contrast to the protein carrying the C-terminal His-tag, the EMRE₁₆₅(M56A) fusion protein with the N-terminal His-tag was clearly observed, indicating that some of the molecules were inserted with their C-termini in the cytoplasm, which, again, was in line with the orientation assay above (Fig. 4A). Possibly, the N-terminal His-tag interferes with proper insertion of the protein in the membrane, which is also supported by the breakdown products observed in the flowthrough and elution fractions.

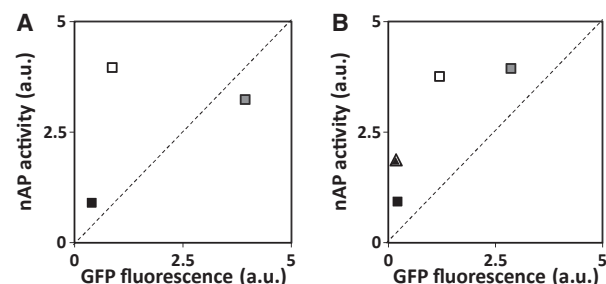


Fig. 4. Membrane orientation of EmrE variants. Normalized AP (nAP) activity is plotted against GFP fluorescence of cells producing EMRE₁₁₀ (gray square), EMRE₁₆₅ (white square) and EMRE₁₆₅(M56A) (black square) with (A) an N-terminal His₆-tag and a C-terminal reporter, and (B) a C-terminal reporter and His₁₀-tag. EMRE₁₆₅(M56A) was induced with 0.04% (black triangle) and 0.1% (white triangle) arabinose.

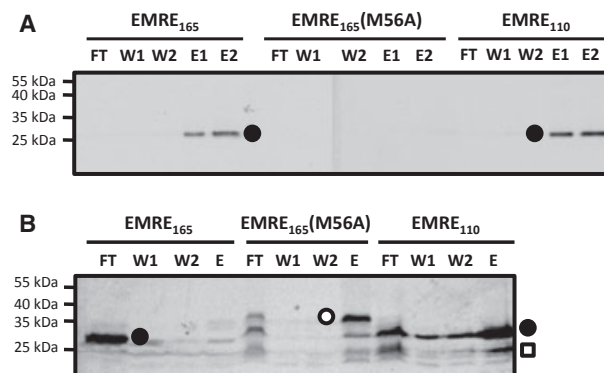


Fig. 5. In-gel fluorescence of GFP fusion proteins isolated from membranes prepared from cells harboring (A) pBADcLIC–GFP vectors encoding the indicated EmrE variants containing a C-terminal His₁₀-tag, and (B) pLIC2 vectors encoding the same EmrE variants containing an N-terminal His₆-tag (Table 2). The EmrE variants were purified by Ni²⁺–nitrilotriacetic acid affinity chromatography, and samples from the flowthrough (FT), wash (W1 and W2) and elution (E1 and E2) steps were analyzed by SDS/PAGE followed by fluorescence imaging of the gel, as detailed in Experimental procedures. ●, EMRE₁₁₀–GFP, ○, EMRE₁₆₅–GFP, □, free GFP.

The results suggest that the long version of EmrE inserts into the membrane in a single orientation, with the C-terminus in the periplasm, which would be in line with the presence of a signal sequence at the N-terminus of the protein (Fig. 3B).

Maturation of EMRE₁₆₅

The genes encoding EMRE₁₁₀ and the EMRE₁₆₅(M56A) were cloned in the pBADhis vector (Table 2) to study the fate of the putative signal sequence present in EMRE₁₆₅. The pBADhis vector produces the inserts with an N-terminal His₆-tag. Following isolation of cytoplasmic membranes from the cells harboring the plasmids, and purification by Ni²⁺-nitrilotriacetic acid affinity chromatography from the solubilized membranes, the expression levels were too low to be detected by staining of the gel after SDS/PAGE (not shown). To enhance the sensitivity of the detection, the biotin acceptor domain (BAD) of the oxaloacetate decarboxylase of *Klebsiella pneumoniae* was inserted in between the His₆-tag and the EmrE protein (pBADBAD vectors; Table 2). BAD is biotinylated *in vivo*, and can be detected with high sensitivity by western blotting with streptactin. BAD fused to the short version of EmrE produced from pBADBAD_EMRE₁₁₀ was readily detected in the membrane fraction as a protein with an apparent molecular mass of ~20–22 kDa, which is in line with the masses of EMRE₁₁₀ and the 10-kDa BAD (Fig. 6). In contrast, no BAD was detected in the membrane fraction from cells expressing EMRE₁₆₅(M56A). Rather, the cytoplasmic fraction of the cells contained a low amount of biotinylated protein with an apparent molecular mass that was slightly more than expected for BAD itself. The results demonstrate that the N-terminal BAD was efficiently removed from the membrane-bound EmrE part.

Activity of EMRE₁₁₀ and EMRE₁₆₅

EmrE makes the cells resistant to toxic compounds. EtBr is a known substrate of EmrE of strain K-12 (EMRE₁₁₀). The ability of the EmrE variants to confer resistance to *E. coli* SF100 was assayed by spotting 10-fold serial dilutions on LB plates containing 500 µg·mL⁻¹ EtBr (Fig. 7). The host *E. coli* SF100 by itself or harboring the pBADhis vector grew well on the plates in the absence of EtBr, whereas growth was completely inhibited in its presence. Plasmid pBADhis_EMRE₁₆₅ producing both EMRE₁₁₀ and EMRE₁₆₅ conferred significant resistance to the cells. Plasmid pBADhis_EMRE₁₆₅(M56A) producing only

EMRE₁₆₅ did not confer resistance, suggesting that EMRE₁₁₀ was responsible for the resistance in the former case. However, surprisingly, plasmid pBADhis_EMRE₁₁₀ producing only EMRE₁₁₀ did not confer resistance. Control experiments showed that cells harboring the three plasmids all showed the same growth on plates without EtBr.

The EmrE proteins produced from the pBADhis vectors carry an N-terminal His₆-tag. The three genes were recloned in pBAD24 vectors that produce the proteins without any tags. The untagged EMRE₁₁₀ made the SF100 cells resistant to EtBr (Fig. 7), demonstrating that the N-terminal His₆-tag inhibited the activity of the protein. The matured EMRE₁₆₅ did not confer resistance, whereas the mixture of the two

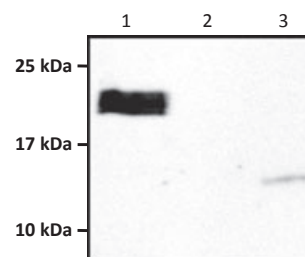


Fig. 6. Processing of EMRE₁₆₅: western blotting of membrane fractions of EMRE₁₁₀ (lane 1), EMRE₁₆₅(M56A) (lane 2), and the cytoplasmic fraction of EMRE₁₆₅(M56A) (lane 3). The proteins were expressed with BAD fused at the N-terminus. Membranes and cytoplasm of cells producing the EmrE variants were separated by centrifugation, after which proteins were separated by SDS/PAGE and, following blotting, biotinylated proteins were detected with streptactin.

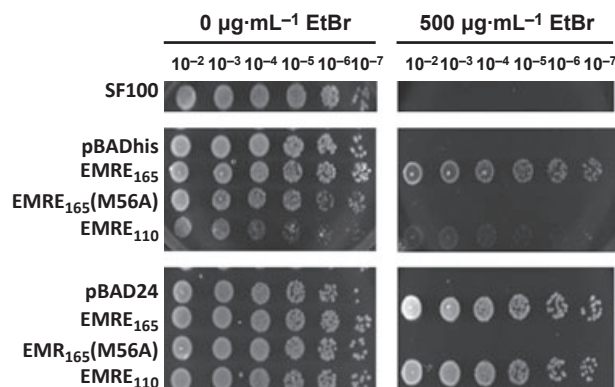


Fig. 7. Activity of EmrE variants. Ten-fold serial dilutions of SF100 cells expressing EMRE₁₆₅, EMRE₁₆₅(M56A) and EMRE₁₁₀ with (top) and without (bottom) an N-terminal His₆-tag were spotted on LB plates supplemented with the indicated concentrations of EtBr.

versions did confer resistance, as was observed with the His-tagged versions. The activity of the latter is explained by the fact that EMRE₁₁₀ produced from pBADhis_EMRE₁₆₅ does not carry a His₆-tag.

The pattern of resistance was the same for the EmrE versions produced from the pLIC1 and pLIC2 vectors when plated on LB plates containing 500 µg·mL⁻¹ EtBr (not shown). Plasmids pLIC1_EMRE₁₆₅ and pLIC2_EMRE₁₆₅ conferred resistance to EtBr, demonstrating, in addition, that EMRE₁₁₀ with the reporters fused at the C-terminus are active proteins. Most importantly, matured EMRE₁₆₅ inserted into the membrane in one orientation did not confer resistance against EtBr to the cells.

Discussion

Evolutionary distribution of the *emrE* gene in *E. coli*

The *emrE* gene is a member of one of a number of gene families that confer resistance to toxic compounds, to allow the organism to survive in hazardous environments. The presence of these genes is required in particular habitats, and is therefore often strain-specific rather than species-specific. The mechanism by which the genes are propagated in a particular species is believed to be through HGT. The analysis of the distribution of the *emrE* gene in 29 *E. coli* strains presented here supports this view, because: (a) different strains contain no, one or two copies of the gene; (b) different genetic contexts of the genes were found; and (c) the phylogenetic tree of the proteins does not correspond to the tree of the strains. Among 29 analyzed strains, eight do not possess the *emrE* gene. Possibly, these particular strains do not encounter the toxic substrates of EmrE in their environment, or the resistance is conferred by another protein. Strains with two *emrE* genes may require a higher capacity to remove toxic compounds because of higher concentrations in their habitats. Analysis of the genetic context of all *emrE* genes revealed two major insertion sites, termed here the FLA and REN sites. The FLA region appears to be more stable than the REN region, which is more variable and contains several genes putatively involved in the process of HGT. The different genetic contexts correlated with the sequence divergence of the *emrE* genes (Fig. 2) and the properties of the encoded proteins, i.e. the long and short versions, suggesting independent evolution of the FLA and REN types. It is not clear why the sequence divergence of the genes correlates with the insertion sites, but not with the phylogeny of the different strains. Apparently, the region

that is horizontally transferred is larger than the *emrE* gene.

Function of EMRE₁₆₅

Two types of *emrE* were found in the genomes of different *E. coli* strains, encoding proteins of different lengths, i.e. EMRE₁₁₀ and EMRE₁₆₅. The ORF encoding the long version contains the ORF encoding the short version, so both EMRE₁₆₅ and EMRE₁₁₀ are likely to be produced. EMRE₁₁₀ is typical of the members of the SMR family. It is a highly hydrophobic, dual topology membrane protein that is active in drug extrusion. In EMRE₁₆₅, this protein is preceded by a putative signal sequence connected by a 34-residue loop. The properties of EMRE₁₆₅ were studied separately by inactivating the translation of the shorter gene encoding EMRE₁₁₀. EMRE₁₆₅(M56A) is inserted into the membrane in one orientation with the C-terminus in the periplasm, and matures through removal of the N-terminal part, which is in agreement with the signal sequence hypothesis. In addition, matured EMRE₁₆₅(M56A) was shown to be inactive in conferring resistance to EtBr, which may be related to the single orientation of the protein in the membrane. EmrE-like proteins are believed to be active as dimers, two types of which can be formed when both orientations of the protein are available in the membrane. One type is built from two monomers with the same orientation (parallel dimer), and the other type from monomers with opposite orientations (antiparallel dimer). Recent studies have shown that antiparallel dimers are more stable than parallel dimers [32], but, in spite of many efforts, it is still not entirely clear which dimers are formed and which is the active configuration. EMRE₁₆₅ inserts into the membrane in a single orientation, and can only form parallel dimers. The lack of activity would support the view that the antiparallel dimer represents the physiological, active dimer complex, but other reasons for the lack of activity, such as the presence of the additional N-terminal region, cannot be excluded. Also, the lack of activity in drug extrusion does not exclude another function for EMRE₁₆₅.

What might be the function of EMRE₁₆₅, which by itself, appears to be inactive? Possibly, an ancestor of EmrE was inserted into the membrane in a single orientation, but, from a physiological point of view, there was an evolutionary advantage in the formation of antiparallel dimers, e.g. to broaden the substrate specificity (reviewed by Bay [33]). In order for this to be possible, part of the protein would need to be inserted into the membrane in the opposite orientation. One

way to achieve this was by manipulation of the insertion process by insertion/deletion of positive charges in the loops, resulting in 'dual topology' proteins. Another possible way would be to produce separately proteins with both orientations via a gene duplication followed by divergence of one of the two genes, yielding the opposite orientation. Both mechanisms can be easily identified in the proteins of the SMR family and other families [11,12,34–36]. A gene reminiscent of the EMRE₁₆₅-encoding gene may represent another solution, yielding two proteins with the opposite orientation, but without requiring a gene duplication event. A single messenger would encode both orientations in two overlapping reading frames. The short ORF would encode the ancestor EmrE that would be forced into the opposite orientation by the signal sequence encoded in the long ORF. Eventually, the solution faded out in evolution, and EMRE₁₆₅ might be an evolutionary relic of this mechanism.

Inhibitory properties of the N-terminal His₆-tag

EmrE is a multidrug transporter that transports a multitude of organic cations. Glu14 positioned in the middle of the membrane in TMS1 is highly conserved in the SMR family, and is believed to be directly involved in substrate and proton binding. It has been suggested that release of two protons into the cytoplasm from the two Glu14 residues in the binding site at the interface of the dimer of two EmrE subunits would allow binding of the positively charged substrate. Subsequently, conformational changes would expose the substrate to the periplasm, and this would be followed by release of the substrate and rebinding of the protons at the periplasmic side. Reorientation of the binding would return the protein to the original state, with the binding site facing the cytoplasm [14,33,37]. The present study has demonstrated that a sequence of six histidines fused at the N-terminus of EMRE₁₁₀ (His₆-tag) interferes with this mechanism. Cells expressing the His-tagged EMRE₁₁₀ could not grow on agar plates containing EtBr. His-tagged EMRE₁₆₅ did not suffer from this inhibition, because of the processing of the protein, which removes the tag, and because of the production of EMRE₁₁₀ without the tag from the same messenger. Possibly, one or more His residues protonated at the imidazole ring bind to the active site and block the activity. The poor affinity of the His⁺ substrate may be well compensated for by the high local concentration. Alternatively, the His-tag at the N-terminus might interfere with dimer formation, explaining the inhibition of EMRE₁₁₀.

Experimental procedures

Strains and growth conditions

E. coli strain SF100 [38] carrying vectors pLIC, pBAD-cLIC, pBAD24 (Invitrogen), pBADhis and pBADBAD (see below) was grown in LB medium supplemented with 50 µg·mL⁻¹ ampicillin at 37 °C under continuous shaking. Overnight cultures were diluted 1 : 30, and grown to a *D*_{660 nm} of 0.6, after which expression was induced with 0.004% arabinose unless otherwise indicated, followed by incubation for another 1.5 h.

DNA manipulations

The constructs used in this study are listed in Table 2. The genes encoding EMRE₁₆₅ and EMRE₁₁₀ were amplified from *E. coli* strains 536 and K-12, respectively, by PCR with *pfu* polymerase (Fermentas, Waltham, MA, USA). Both genes were cloned into pLIC1 and pLIC2 plasmids by ligase-independent cloning (LIC), with primers that contained an overhang sequence compatible with the pLIC cassette, as described previously [13,31]. In order to create single-stranded overhangs, PCR products were treated for 30 min at room temperature with T4 DNA polymerase (Fermentas) in the presence of dCTP. Similarly, plasmids pLIC1 and pLIC2 were treated with T4 DNA polymerase in the presence of dGTP after linearization with *Swa*I (Fermentas). T4 polymerase was inactivated by incubation at 75 °C for 20 min. The T4 polymerase-treated PCR products and vectors were combined, and, after 5 min of incubation at room temperature, transformed into *E. coli* SF100. The start codons of EMRE₁₁₀ were inactivated (M56A mutants) by PCR of the whole plasmids with a forward primer containing the mutation. PCR products were purified, treated with *Dpn*I, and transformed into *E. coli* SF100. Gene expression from pLIC1 and pLIC2 results in transcriptional fusion of the cloned genes with AP and GFP, respectively. The pLIC1 and pLIC2 vectors encode EmrE variants with an N-terminal His₆-tag linked by a Leu to the protein, and, at the C-terminus, the linker sequence QNSGVVP followed by the reporter protein.

Cloning into the LIC vector pBADcLIC-GFP (a kind gift from E. Geertsma, University of Zurich, Switzerland) was performed as described previously [39]. The primers used in cloning contained an overhang sequence compatible with the cLIC cassette (forward, 5'-ATGGGTGGTGGATTGCT-3'; reverse, 5'-TTGGAAGTATAAATTTT C-3'). Plasmid pBADcLIC-AP was derived from pBAD-cLIC-GFP. The gene encoding AP was amplified from pLIC1 with a forward primer containing an *Swa*I restriction site and a reverse primer containing an *Spe*I restriction site. The GFP gene was restricted from pBADcLIC-GFP with the same enzymes, and replaced with the AP gene encoding PCR product after digestion with *Swa*I and *Spe*I.

Insertion of the EmrE-encoding genes was performed as described above for pBADcLIC–GFP. Expression of the genes from pBADcLIC plasmids results in transcriptional fusions with a tobacco etch virus cleavage site, the reporter protein, and a His₁₀-tag.

Cloning into pBAD24 (Invitrogen) and its derivative pBADhis [40] was performed by amplification of the *emrE* genes with forward primers containing an *Nco*I site and a reverse primer containing an *Xba*I site. Plasmids and PCR products were digested with *Nco*I and *Xba*I, and ligated with T4 DNA ligase (Fermentas) by incubation at room temperature for 1 h.

Finally, the gene encoding BAD was restricted from pBADCitS [41] with *Nco*I, and inserted into plasmids pBADhis_EMRE₁₁₀ and pBADhis_EMRE₁₆₅(M56A) containing the different *emrE* variants. The expressed proteins contained a His₆-tag followed by BAD at their N-termini.

GFP and AP activity of whole cells

GFP fluorescence

Cells from 2 mL of culture were washed once, and resuspended in 50 mM Tris/HCl (pH 8.0), 200 mM NaCl and 15 mM EDTA to a $D_{660\text{ nm}}$ of 0.2. *N*-dodecyl β -D-maltoside was added to the suspension to a final concentration of 0.5% (w/v). The fluorescence was measured with an Aminco-Bowman Series 2 Spectrometer at an excitation wavelength of 468 nm and an emission wavelength of 507 nm. For each sample, background fluorescence of cells harboring the vector without insert and reporter was subtracted. Experiments were performed in triplicate.

AP activity

Cells from 2 mL of culture were washed once and resuspended in 1 mL of 1 M Tris/HCl (pH 8.0), and the $D_{660\text{ nm}}$ was measured. Following equilibration of 500 μ L of the suspension at 37 °C for 5 min, *p*-nitrophenyl phosphate (Sigma Aldrich, Zwijndrecht, The Netherlands) was added to a final concentration of 1.4 mg·mL^{−1}. The reaction was stopped with 1 M K₂HPO₄ when a yellow color developed. AP activity was expressed in Miller units [42]. Measurements were performed in triplicate. GFP fluorescence and AP activities were normalized by the ratio of the averages of the positive PhoA and GFP activities of a set of 27 SMR proteins, as described previously [13,31].

Resistance assay

Cells were grown as described above. After induction with 0.004% arabinose, cells were allowed to grow for another hour. After this time, all cultures were diluted to a $D_{660\text{ nm}}$ of 0.6, and 10-fold serial dilutions were prepared. Five microliters of each dilution in the range of 10^{−2} to 10^{−7}

was spotted onto plates containing 500 μ g·mL^{−1} EtBr. The plates contained no ampicillin. Plates were incubated overnight at 37 °C.

Protein purification

His-tagged proteins were purified by Ni²⁺–nitrilotriacetic acid affinity chromatography with a small-scale purification protocol, as described previously [43]. *E. coli* SF100 cells from 200-mL cultures were washed twice with 50 mM potassium phosphate (pH 7.0), and resuspended in 2 mL of the same buffer. Cells were broken by sonication (Soni-prep 150, MSE Ltd, London, UK) with nine cycles of 15 s on and 45 s off while the suspension was kept on ice. Unbroken cells and debris were removed by low-speed centrifugation (10 min, 7600 *g*). The cytoplasmic fraction was incubated with Ni²⁺–nitrilotriacetic acid. Membranes were collected by ultracentrifugation with a Beckman TLA 100.4 rotor (25 min, 347 000 *g* at 4 °C), and washed once with 50 mM potassium phosphate (pH 7.0). Subsequently, the membranes were solubilized for 30 min on ice in 50 mM potassium phosphate (pH 8.0), 400 mM NaCl and 10% glycerol containing 1% Triton X-100. Undissolved material was removed by ultracentrifugation at 80 000 r.p.m. for 25 min. The supernatant was mixed with 100 μ L of Ni²⁺–nitrilotriacetic acid resin, and incubated overnight. The column material was pelleted by pulse centrifugation with a table-top centrifuge. The supernatant was discarded. The resin was washed twice with 500 μ L of 50 mM potassium phosphate (pH 8.0), 300 mM NaCl and 10% glycerol containing 40 mM imidazole and 0.1% Triton X-100. Bound proteins were eluted with 50 μ L of 50 mM potassium phosphate (pH 8.0), 300 mM NaCl and 10% glycerol containing 150 mM imidazole and 0.1% Triton X-100 and, when indicated, with 50 μ L of the same buffer containing 300 mM imidazole. Samples from the different purification steps were run on a 15% SDS/PAGE gel, after which GFP-tagged proteins were visualized by fluorescence imaging of the gel with a Fuji LAS-4000 imager. BAD-tagged proteins were detected by western blotting. The proteins were transferred to a poly(vinylidene difluoride) membrane with the semidry transfer technique. The membrane was incubated overnight at 4 °C with NaCl/P_i (with 0.1% Tween-20) supplemented with 0.2% I-block (Sigma Aldrich), and this was followed by three washes for 10 min each with NaCl/P_i-T containing 0.1% I-block, incubation for 1.5 h with Strept-actin (IBA, Göttingen, Germany), and finally three washes for 10 min each. The chemiluminescence signal was obtained with CDP-Star (Roche), and visualized with a Fuji LAS-4000 imager.

References

- Doi Y, Shibata N, Shibayama K, Kamachi K, Kurokawa H, Yokoyama K, Yagi T & Arakawa Y

- (2002) Characterization of a novel plasmid-mediated cephalosporinase (CMY-9) and its genetic environment in an *Escherichia coli* clinical isolate. *Antimicrob Agents Chemother* **46**, 2427–2434.
- 2 Miriagou V, Tzouvelekis LS, Lebessi E, Vatopoulos AC, Carattoli A & Tzelepi E (2004) CMY-13, a novel inducible cephalosporinase encoded by an *Escherichia coli* plasmid. *Antimicrob Agents Chemother* **48**, 3172–3174.
 - 3 Jęlaszewicz J, Młynarczyk G & Młynarczyk A (2000) Antibiotic resistance in Gram-positive cocci. *Int J Antimicrob Agents* **16**, 473–478.
 - 4 Li XZ, Poole K & Nikaido H (2003) Contributions of MexAB–OprM and an EmrE homolog in intrinsic resistance of *Pseudomonas aeruginosa* to aminoglycosides and dyes. *Antimicrob Agents Chemother* **47**, 27–33.
 - 5 Sidhu MS, Heir E, Sorum H & Holck A (2001) Genetic linkage between resistance to quaternary ammonium compounds and beta-lactam antibiotics in food-related *Staphylococcus* spp. *Microb Drug Resist* **7**, 363–371.
 - 6 Sidhu MS, Heir E, Leegaard T, Wiger K & Holck A (2002) Frequency of disinfectant resistance genes and genetic linkage with beta-lactamase transposon Tn552 among clinical staphylococci. *Antimicrob Agents Chemother* **46**, 2797–2803.
 - 7 Vourli S, Tzouvelekis LS, Tzelepi E, Lebessi E, Legakis NJ & Miriagou V (2003) Characterization of In111, a class I integron that carries the extended-spectrum beta-lactamase gene bla_{IBC-1}. *FEMS Microbiol Lett* **225**, 149–153.
 - 8 Yerushalmi H, Lebendiker M & Schuldiner S (1995) EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H⁺ and is soluble in organic solvents. *J Biol Chem* **270**, 6856–6863.
 - 9 Paulsen IT, Skurray RA, Tam R, Saier MH Jr, Turner RJ, Weiner JH, Goldberg EB & Grinius LL (1996) The SMR family: a novel family of multidrug efflux proteins involved with the efflux of lipophilic drugs. *Mol Microbiol* **19**, 1167–1175.
 - 10 Daley DO, Rapp M, Granseth E, Mele K, Drew D & von Heijne G (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308**, 1321–1323.
 - 11 Rapp M, Granseth E, Seppälä S & von Heijne G (2006) Identification and evolution of dual-topology membrane proteins. *Nat Struct Mol Biol* **13**, 112–116.
 - 12 Rapp M, Granseth E, Seppälä S & von Heijne G (2007) Emulating membrane protein evolution by rational design. *Science* **315**, 1282–1284.
 - 13 Kolbusz MA, ter Horst R, Slotboom DJ & Lolkema JS (2010) Orientation of small multidrug resistance transporter subunits in the membrane: correlation with the positive-inside rule. *J Mol Biol* **402**, 127–138.
 - 14 Ubarretxena-Belandia I, Baldwin JM, Schuldiner S & Tate CG (2003) Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J* **22**, 6175–6181.
 - 15 Ubarretxena-Belandia I & Tate CG (2004) New insights into the structure and oligomeric state of the bacterial multidrug transporter EmrE: an unusual asymmetric homo-dimer. *J Mol Biol* **340**, 797–808.
 - 16 Chen YJ, Pornillos O, Lieu S, Ma C, Chen AP & Chang G (2007) X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci USA* **104**, 18999–19004.
 - 17 Toshifumi N, Kouyama T, Kurata Y, Kikukawa T, Miyauchi S & Kamo N (2007) Anti-parallel membrane topology of a homo-dimeric multidrug transporter, EmrE. *J Biochem* **142**, 621–625.
 - 18 Korkhov VM & Tate GM (2009) An emerging consensus for the structure of EmrE. *Acta Crystallogr D* **65**, 186–192.
 - 19 Schuldiner S (2007a) When biochemistry meets structural biology: the cautionary tale of EmrE. *Trends Biochem Sci* **32**, 252–258. Erratum in *Trends Biochem Sci* **32**, 300.
 - 20 Schuldiner S (2007b) Controversy over EmrE structure. *Science* **317**, 748–751.
 - 21 Nasie I, Steiner-Mordoch S, Gold A & Schuldiner S (2010) Topologically random insertion of EmrE supports a pathway for evolution of inverted repeats in ion-coupled transporter. *J Biol Chem* **285**, 15234–15244.
 - 22 Schuldiner S (2012) Undecided membrane proteins insert in random topologies. Up, down and sideways: it does not really matter. *Trends Biochem Sci* **37**, 215–219.
 - 23 Sims GE & Kim SH (2011) Whole-genome phylogeny of *Escherichia coli*/Shigella group by feature frequency profiles (FFPs). *Proc Natl Acad Sci USA* **108**, 8329–8334.
 - 24 Krogh A, Larsson B, von Heijne G & Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.
 - 25 Arkin IT, Russ WP, Lebendiker M & Schuldiner S (1996) Determining a secondary structure and orientation of EmrE, a multi-drug transporter, indicates a transmembrane four-helix bundle. *Biochemistry* **35**, 7233–7238.
 - 26 Schwaiger M, Lebendiker M, Yerushalmi H, Coles M, Gröger A, Schwarz C, Schuldiner S & Kessler H (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein. *Eur J Biochem* **254**, 610–619.
 - 27 Elbaz Y, Salomon T & Schuldiner S (2008) Identification of a glycine motif required for packing in EmrE, a multidrug transporter from *Escherichia coli*. *J Biol Chem* **283**, 12276–12283.
 - 28 Schuldiner S (2009) EMRE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim Biophys Acta* **1794**, 748–762.

- 29 Nielsen H, Engelbrecht J, Brunak S & von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**, 1–6.
- 30 Bendtsen JD, Nielsen H, von Heijne G & Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**, 783–795.
- 31 ter Horst R & Lolkema JS (2010) Rapid screening of membrane topology of secondary transport proteins. *Biochim Biophys Acta* **1798**, 672–680.
- 32 Lloris-Garcera P, Bianchi F, Slusky JS, Seppälä S, Daley DO & von Heijne G (2012) Antiparallel dimers of the small multidrug-resistance protein EmrE are more stable than parallel dimers. *J Biol Chem* **287**, 26052–26059.
- 33 Bay DC, Rommens KL & Turner RJ (2007) Small multidrug resistance family: a multidrug transporter family that continues to grow. *Biochim Biophys Acta* **1778**, 1814–1838.
- 34 Pornillos O & Chang G (2006) Inverted repeat domains in membrane proteins. *FEBS Lett* **580**, 358–362.
- 35 Lolkema JS, Dobrowolski A & Slotboom DJ (2008) Evolution of antiparallel two-domain membrane proteins: tracing multiple gene duplication events in the DUF606 family. *J Mol Biol* **378**, 596–606.
- 36 Bay DC & Turner RJ (2009) Diversity and evolution of the small multidrug resistance protein family. *BMC Evol Biol* **9**, 140. doi:10.1186/1471-2148-9-140.
- 37 Fleishman SJ, Harrington SE, Enosh A, Halperin D, Tate C & Ben-Tal N (2006) Quasi-symmetry in the cryo-EM structure of EmrE provides the key to modeling its transmembrane domain. *J Mol Biol* **364**, 54–67.
- 38 Baneyx F & Georgiou G (1990) In vivo degradation of secreted fusion proteins by the Escherichia coli outer membrane protease OmpT. *J Bacteriol* **172**, 491–494.
- 39 Geertsma E & Poolman B (2007) High-throughput cloning and expression in recalcitrant bacteria. *Nat Methods* **4**, 705–707.
- 40 Sobczak I & Lolkema JS (2003) Accessibility of cysteine residues in a cytoplasmic loop of CitS of *Klebsiella pneumoniae* is controlled by the catalytic state of the transporter. *Biochemistry* **42**, 9789–9796.
- 41 Krupnik T, Dobrowolski A & Lolkema JS (2011) Cross-linking of dimeric CitS and GltS transport proteins. *Mol Membr Biol* **28**, 243–253.
- 42 Manoil C (1991) Analysis of membrane protein topology using alkaline phosphatase and beta-galactosidase gene fusions. *Methods Cell Biol* **34**, 61–75.
- 43 Dobrowolski A, Sobczak-Elbourne I & Lolkema JS (2007) Membrane topology prediction by hydropathy profile alignment: membrane topology of the Na⁺-glutamate transporter GltS. *Biochemistry* **46**, 2326–2332.